

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**ScienceDirect**

Transportation Research Procedia 3 (2014) 942 – 951

---

---

**Transportation  
Research  
Procedia**

---

---

[www.elsevier.com/locate/procedia](http://www.elsevier.com/locate/procedia)

17th Meeting of the EURO Working Group on Transportation, EWGT2014, 2-4 July 2014,  
Sevilla, Spain

## Real-time single detector vehicle classification

Joel Dodsworth\*, Simon Shepherd, Ronghui Liu

Institute for Transport Studies, University of Leeds, 34-40 University Road, Leeds, LS2 9JT, United Kingdom

---

### Abstract

The estimation of speed from a single vehicle detector has been a popular area of research largely due to the potential for reduced maintenance and installation costs associated with the more conventional method of using two closely spaced detectors with a known offset distance. Speed cannot be directly measured using a single detector but flow and occupancy measurements (the percentage of time a detector is occupied in any given sample) can be recorded and used to estimate the speed and subsequently, using the vehicle presence time, the length of vehicles.

The approximation of space occupancy from the time occupancy provided by a single short detector requires that vehicle speeds in each sampling interval are constant and subsequently research on this subject is concentrated on freeway applications where speeds can generally be assumed constant during each sampling interval. However, in urban situations and particularly in the vicinity of traffic signal controlled junctions it cannot be assumed that this is the case due to the rapid variation in vehicle speed. Centralised on-line adaptive traffic signal optimisation strategies use data received from detectors placed on approaches upstream of junctions to feed information into an underlying traffic model. Despite the advancement in communication technology 4Hz and 10Hz protocols are still used widely in Urban Traffic Control (UTC) and consequently this research explores the accuracy of speed estimation and basic vehicle length classification that can be achieved using data sampled at these rates.

For this application data must be available within a time window of a few seconds to be useful in the optimisation process. This work investigates what level of accuracy can be obtained by estimating vehicle speed and length immediately following a vehicle leaving the detection zone using a process of speed matching to provide useful data to an on-line traffic model.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014

**Keywords:** Speed estimation; vehicle classification; single detector; traffic signal optimisation; urban links

---

\* Tel.: +44-113-343-5325.

E-mail address: [tsjad@leeds.ac.uk](mailto:tsjad@leeds.ac.uk)

## 1. Introduction

Traffic surveillance techniques have been used for decades to monitor traffic conditions on the road network. The primary purpose of monitoring traffic conditions is to enable traffic control strategies to be implemented either manually or automatically to maximise efficiency or to achieve some other objective according to local policy. Traditionally inductive loop detectors buried in the surface of the road have been employed to detect the presence of vehicles and are still the most common method of traffic surveillance despite the emergence of other technology including above ground detection.

In urban road networks traffic signals are often used to manage traffic with larger networks often controlled centrally using a UTC (Urban Traffic Control) system. The development of more intelligent adaptive traffic signal control strategies led to the requirement for detector data to be transmitted to a central system frequently enough to be used in a real-time optimisation process. Limitations in communication bandwidth and computational power during the early development years of UTC led to the use of second-by-second data transfer between local traffic signal controllers and the central UTC system.

One of the most prevalent on-line traffic signal optimisation tools is SCOOT (Split Cycle Offset Optimisation Technique) originally developed by TRRL in the 1970s (Robertson and Bretherton, 1991). SCOOT makes use of detector data received from detectors positioned across the network on each approach to traffic signals. Data is actually received as four quarter-second bits per second and converted into LPU's (Link Profile Units); used to convert the presence information into PCU (Passenger Car Unit) values for use in the SCOOT traffic model.

Today computational power has increased dramatically and the advent of the internet and associated investment in infrastructure has greatly improved the bandwidth available for communication. However, the most common communication protocols for centralised signal control still transfer low resolution vehicle presence data. This work investigates whether any more detailed information can be inferred from the existing protocols to improve the traffic models used in the traffic signal optimisation process and subsequently discusses how the results can inform future work and development.

The research is part of a wider attempt to develop a more sophisticated traffic model for traffic signal optimisation that intends to make use of vehicle classification to better model the effects of vehicle mix on queue discharge time and platoon progression as well as providing more accurate spatial queue information. The research investigates whether existing methods used for estimating freeway vehicle speed from a single detector can be adapted for urban conditions and what accuracy of basic vehicle length classification can be expected. Speed estimation in this case is a means to subsequently classify vehicles by length and is not the main objective of the work although results could prove useful for other applications.

<b>Nomenclature</b>		
EVL	effective vehicle length	(m)
FT	factor threshold	
g	reciprocal of MEVL	(m <sup>-1</sup> )
HOT	high occupancy threshold	(%)
L <sub>det</sub>	effective detector length	(m)
L <sub>veh</sub>	vehicle length	(m)
LOT	low occupancy threshold	(%)
LV	long vehicle	
MEVL	mean effective vehicle length	(m)
MSI	maximum permissible speed increase	(ms <sup>-1</sup> )
n	number of vehicles in a specified sample interval	(veh)

O	occupancy in a specified sample interval ( $\sum \tau / T$ )	(%)
O <sub>n</sub>	occupancy in a sample of size n vehicles	(%)
OTF	factor of $\tau$ between current and previous vehicle	
OTFI	factor of $\tau$ between previous and current vehicle	
SV	short vehicle	
$\tau$	vehicle detector ‘on-time’	(s)
T <sub>i</sub>	initial sample interval length	(s)
T <sub>p</sub>	proposed sample interval length (rear to rear headway)	(s)
v <sub>lb</sub>	lower bound speed (used in algorithm)	(ms <sup>-1</sup> )
v <sub>lv</sub>	estimated long vehicle speed (used in algorithm)	(ms <sup>-1</sup> )
v <sub>oub</sub>	high occupancy limited upper bound speed (used in algorithm)	(ms <sup>-1</sup> )
v <sub>pv</sub>	previous vehicle speed estimate (used in algorithm)	(ms <sup>-1</sup> )
v <sub>sv</sub>	estimated short vehicle speed (used in algorithm)	(ms <sup>-1</sup> )
$\bar{v}$	space-mean speed	(ms <sup>-1</sup> )

### 1.1. Background

Speed estimation from a single detector has been the subject of many research papers with complex filtering methods often being proposed in an attempt to improve accuracy. There are two principal methods of estimating speed from a single detector, one of which makes use of the signature profile created by a vehicle chassis passing through the electromagnetic field of an inductive loop (or a magnetic sensor) to classify vehicles. Sun and Ritchie (1999) is an example of this technique and various commercial products have been developed (RTEM, 2013) that can provide redundancy at dual loop ATC (Automatic Traffic Counter) sites in case one of the loops fails. The other method used to estimate speed draws on the relationship between flow and occupancy to calculate a space mean speed over a specified time interval using processed vehicle presence data. This work focusses on the latter method as in theory it can be used with various types of detector, doesn't require any additional equipment in the roadside cabinet to process data and can use existing communication protocols.

$$\bar{v} = \frac{n}{T_i \times O \times g} \quad (1)$$

$$MEVL = \frac{\sum (L_{veh} + L_{det})}{n} = \frac{1}{g} \quad (2)$$

Fundamental to the method used for this research is the relationship between speed, flow and occupancy shown by Eq. (1) the derivation of which is described by Kwon et al. (2003). This relationship enables the speed of a sample of vehicles to be estimated on the assumption that vehicle speeds within the measured sample interval are constant. Both components of the MEVL can be estimated but one, the vehicle length, is not constant.

In the case of inductive loops there are numerous factors that can contribute to  $L_{det}$  varying from one detector to another even if the design length of each loop is identical. These include variations in the buried depth of the cable, the length of the cable run from the detector to the roadside cabinet and the sensitivity of the monitoring equipment. However, at a specific site this value can be assumed constant as the value does not change significantly over time. The length of each vehicle is more difficult to calibrate as it cannot be measured directly from the detector and is therefore usually given as an average from observed data for the associated link. This value of the so called  $g$ -factor has been subject to various research including Wang and Nihan (2003), Coifman (2001) and Zhanfeng et al. (2001) with its value subject to extensive debate.

It is clear that the  $g$ -factor is not in fact constant as it depends on the vehicle mix in any measured time interval which can vary significantly from a mean length measured over a longer time period. As the length of a sample interval increases it becomes more likely a representative vehicle mix will be captured but it also increases the likelihood that the actual vehicle speed during the interval is not constant, particularly in saturated conditions, thus severely reducing the effectiveness of the speed estimate.

Some of the previous work on this subject has attempted to estimate the  $g$ -factor dynamically in an effort to address the issues associated with vehicle mix. For example, Coifman (2001) proposed that the  $g$ -factor is calculated during freeflow conditions using a threshold of occupancy below which traffic is assumed to be freeflowing. The calculated  $g$ -factor is then used during the more congested periods to estimate vehicle speed. Another method proposed by Wang and Nihan (2003) fixes the  $g$ -factor using the MEVL for short vehicles. Measurements for intervals that include long vehicles are then screened out to enable the speed to be estimated. The estimated speed is subsequently used to calculate the MEVL for each interval and further processing is then undertaken to identify vehicle composition.

The recurring feature of the above methods is that they all estimate speed over a specific time interval ranging between 20-30 seconds to, in some cases, minutes in length or in the case of Coifman and Kim (2009), require a minimum sample size of vehicles (33). The vast majority of algorithms have been produced to estimate freeway speed where it is likely that for the majority of the time conditions will be freeflow and as a result the assumption that speeds are constant during each sample interval holds.

The methods discussed are useful for their applications and pose no problem if the data is being collected offline for later analysis or even if the data is to be used to detect congestion or incidents. However, for the purpose of this work it is necessary that each vehicle is classified instantly upon leaving the detector so that the information can be fed into an on-line traffic model. In addition, the proposed use of the speed estimation algorithms developed in this work is for traffic signal controlled junctions where speeds are likely to fluctuate significantly over short periods of time, particularly in oversaturated conditions. This fluctuation in speed significantly reduces the accuracy of methods that use longer sampling periods as they assume speed is constant for the duration of each sampling interval.

## 1.2. Investigation

Early testing for this work showed that using fixed sample intervals produced a higher error rate than using a fixed sample size with a varying time interval. However, there are still problems associated with larger sample sizes due to the averaging effect that occurs. The results of speed estimation from a single detector using Eq. (1) with three sample sizes are plotted in Fig. 1 against the actual vehicle speeds at that point recorded from a micro-simulation model. As can be observed from the plot there is a significant lag in the estimation process for the larger sample sizes that make them unsuitable for the application of this work. The single vehicle sample follows the rapid change in speed much more closely but conversely, at freeflow speed, it is much more erratic when compared with the larger sample size.

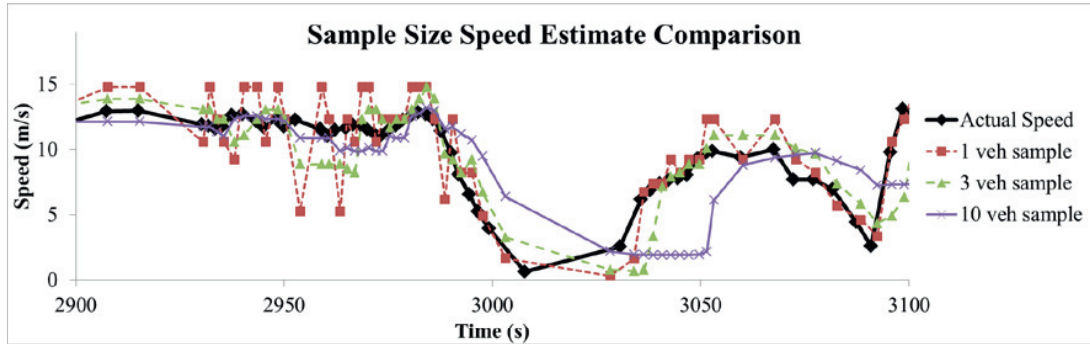


Fig. 1. An example of different sample size speed estimation errors at freeflow and during queue conditions.

Evidently the use of a single vehicle sample increases the magnitude of error resulting from varying vehicle lengths within a sample as there is no longer a mean vehicle length. The downside is that if the EVL is based on the length of a short vehicle but the detected vehicle is in fact a long vehicle the speed could be underestimated by a factor of 3 to 4 (hence the erratic nature of the estimation shown in Fig. 1). In reality the advantage of following the rapidly varying speed more closely outweighs the unstable nature of the single vehicle sample at freeflow speed and that problem can be largely overcome by employing some techniques to reduce error.

For a single vehicle sample size  $n = 1$  and as such the MEVL becomes EVL. Additionally, by substituting Eq. (2) into Eq. (1) it is possible to reduce the equation to a simpler form, see Eq. (3).

$$\bar{v} = \frac{EVL}{\tau} \quad (3)$$

## 2. Methodology

In order to develop the speed estimation algorithm a simulation platform was set up using the PTV-Vissim micro-simulation software package. A simple traffic signal controlled junction was configured in Vissim with a two metre long detector placed 150m from the stopline, see Fig. 2. The simulations were run with a resolution of 10 time steps per simulation second and the vehicle presence data from the detector processed at 4Hz and 10Hz by an external software module developed for this work. The external module processed the presence information to provide flow, occupancy and  $\tau$  for spreadsheet analysis. In order to assess the accuracy of the speed estimation algorithm the simulation vehicle record information was also output from Vissim and filtered to show the actual speed, length and time of each vehicle crossing the detector.

The requirement of the simulation was to represent the traffic conditions experienced in an urban location where single detectors are often used for traffic signal control strategies. In this test the traffic signals in the micro-simulation model were configured to operate a fixed-time cycle time with traffic demand on the approach link varied to create periods of under-saturation and oversaturation. The desired vehicle speeds used in the initial simulation were based on observed data collected by a permanent ATC site in Rotherham, South Yorkshire, UK at a location with a 40mph speed limit.

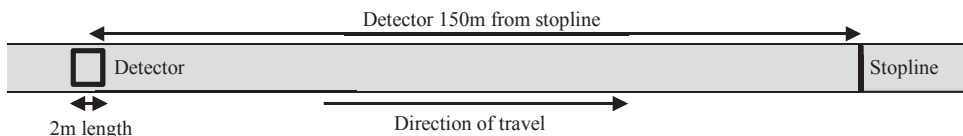


Fig. 2. Model schematic.

$$LV\text{ ACCURACY} = \frac{\sum(LV_{correct})}{\sum(LV_{correct} + LV_{missed} + LV_{wrong})} \quad (4)$$

Performance of the algorithm was measured by calculating the percentage of correctly classified LVs, see Eq. (4), over five simulation runs with varying random seeds for under-saturated and oversaturated conditions. The overall percentage of correctly classified vehicles could also be measured to provide another gauge of the effectiveness of the algorithm but this is significantly biased by the vehicle mix (i.e. a smaller distribution of vehicles across classes will result in an expected decrease in classification error).

### 2.1. Potential sources of error

For this work the detector information was collected at 4Hz to mimic the data received by SCOOT and at 10Hz in order to assess the increase in accuracy that can be expected from the improved data resolution. There is an inherent measurement error resulting from the rounding of the analog vehicle detection profile when it is processed into a digital signal that increases as the data resolution decreases. For example, an average car could be expected to spend approximately 0.5s on a detector when travelling at  $15\text{ms}^{-1}$ , 0.62s at  $12\text{ms}^{-1}$  and 0.39s at  $19\text{ms}^{-1}$ . At 4Hz these could all be rounded to 0.5s resulting in significant speed estimation error. Increasing the resolution of the vehicle presence sampling reduces the influence of this error as can be observed in Fig. 3. in which measured values of  $\tau$  from the same simulation model using 4Hz and 10Hz resolution processing are compared.

Additionally, the method of packaging the data into quarter second intervals may vary between equipment manufacturer and depending on the age of the hardware. For example, newer hardware is capable of recording a much higher resolution of data and can therefore store data at resolutions that can be more easily downsampled to 4Hz but even the way it is downsampled can vary (Peek, 2012). It is clear from this that there can be a significant degree of uncertainty in the information that a centralised system receives from detectors on street.

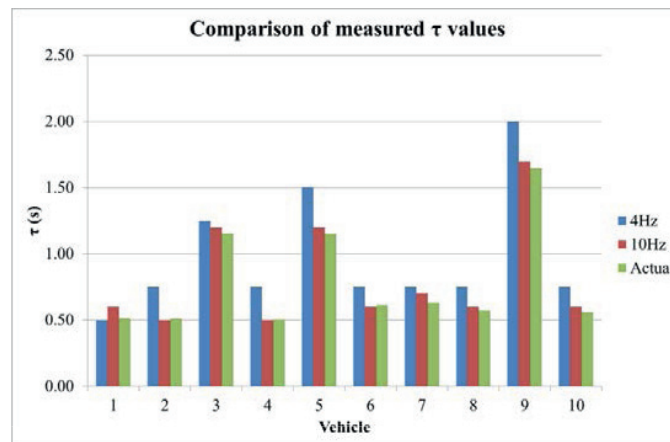


Fig. 3. Example of measurement error at different data resolutions.

## 3. Algorithm Development

The initial speed estimates are calculated using Eq. (3). The value of  $L_{\text{det}}$  varied in testing for 4Hz and 10Hz resolution data, being approximately 3.5m and 2m respectively. The difference in the two values can be explained by the rounding errors detailed previously that result in consistent overestimations of  $\tau$  in the 4Hz scenario.

Speed estimates are calculated at the beginning of each run through the algorithm based on nominal short and long vehicle lengths, in this case an  $L_{\text{veh}}$  of 4.5m and 16m. The short vehicle estimate is treated as the default speed estimate and is then subjected to a series of filters that adjust the initial speed estimate based on various conditions and, if certain criteria are met, either discard the short vehicle speed estimate in favour of the long vehicle estimate or confirm that the short vehicle speed estimate should indeed be used, see Fig. 5.

### 3.1. Low and high occupancy speed correction

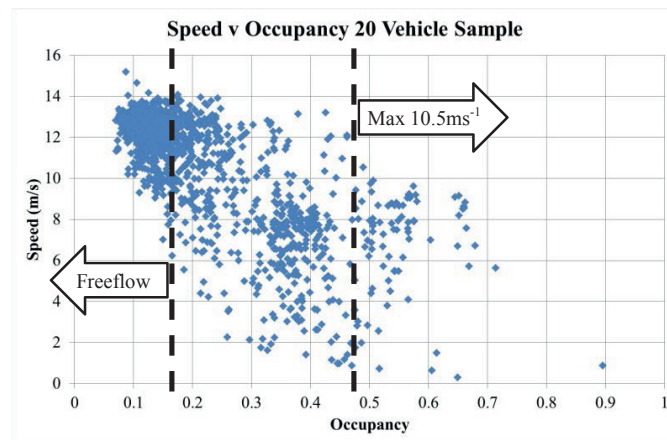
Coifman (2001) proposed that for detector occupancy values below a specific threshold an average freeflow speed should be used to improve estimation accuracy. This work further explores the idea that assumptions can be made based on observed data that improve the estimation of speed and subsequently vehicle length. For the low and high occupancy filtering a larger sample size of 20 vehicles was used as it performed best during initial testing.

Results from simulation runs were used to select a low occupancy threshold value, in this case 16% occupancy, see Fig. 4. Rather than assuming that all vehicles below this occupancy were travelling at an average freeflow speed the filter only adjusts the initial speed estimate if it is less than a freeflow lower bound speed, in this case  $2.5\text{ms}^{-1}$  below the observed average freeflow speed. Likewise initial speed estimates above a freeflow upper bound speed, in this case  $2.5\text{ms}^{-1}$  above the observed average freeflow speed, were adjusted to the upper bound speed.

Using the same method for high occupancy values the results from simulation runs were used to select a high occupancy threshold value, above which vehicles are assumed to travel no faster than a specified speed, in this case  $10.5\text{ms}^{-1}$ .

### 3.2. Low speed

At very low speed the estimation process becomes ineffective due to the difference in speed from a single vehicle entering to leaving the detection zone. For this reason the speed is selected on a purely statistical basis depending on which is the dominant vehicle class using Eq. (3) with the EVL calculated using the dominant vehicle class length. It is clear that this will affect accuracy more significantly if there is a high proportion of LVs. For clarity this is not shown in Fig. 5.



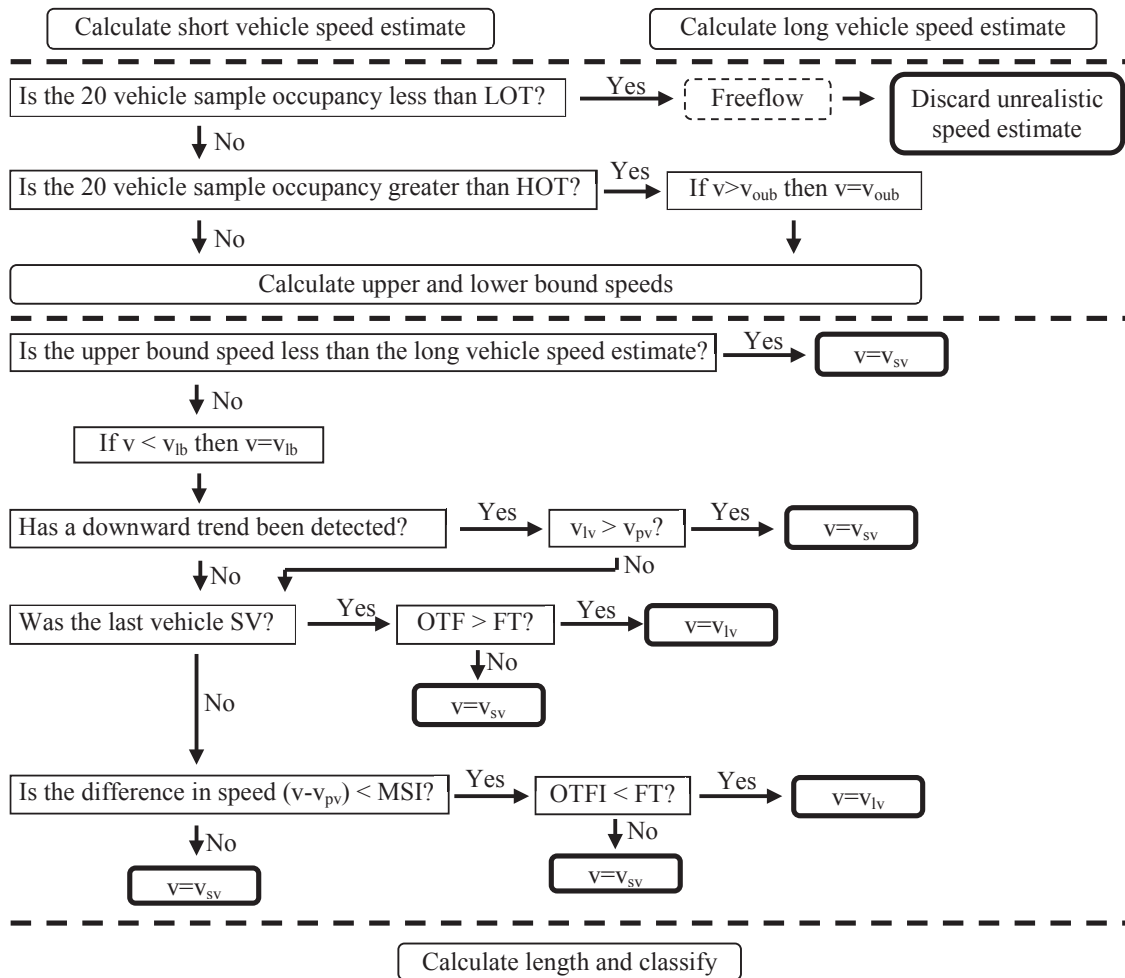


Fig. 5. Speed estimation and classification algorithm schematic.

### 3.3. Upper and lower bound speeds

The algorithm is primarily based on the confidence of the previous vehicle classification. For each vehicle upper and lower bounds are set for the speed estimate based on the confidence of the previous vehicle estimate and the gap between vehicles. These are calculated using a specified maximum gradient value taken from observed data to limit the possible speed difference. If there is a high LV proportion then confidence in previous speed estimates will be reduced as there are fewer opportunities to rule out LV classification on the basis of the initial LV speed estimate. Similarly, larger gaps between vehicles will reduce the confidence of the current vehicle speed estimate.

### 3.4. Downward trend

When the queue from the traffic signals extends far enough it begins to affect approaching vehicle speeds over the detector. It is very rare for consecutive vehicle speeds to increase until the queue begins to discharge. Where a downward trend is detected by consecutive significant drops in vehicle speed a flag is set that prevents the current vehicle speed estimate from being greater than the previous estimate.



### 3.5. Classification based speed estimation

The final step of the algorithm uses the factor of  $\tau$  between the current and preceding vehicles see Fig. 5. If the previous vehicle was classified as LV then unless the current vehicle  $\tau$  is less than ~65% of the previous vehicle  $\tau$  the speed from the last algorithm step is not altered. Alternatively, if the previous vehicle was classified SV then a maximum permissible speed increase from the preceding speed estimate is first applied. Assuming that criterion is met then the current vehicle  $\tau$  must be greater than 150% of the previous vehicle  $\tau$  for the speed to be altered.

### 3.6. Classification

Classification of the vehicles is performed simply by multiplying the speed produced by the algorithm with the vehicle  $\tau$  and subtracting  $L_{det}$ . The vehicle is then classed as LV if the estimated length is greater than the LV threshold length (in this case 14m as articulated HGVs are approximately 16m in length).

## 4. Results

The results in Table 1 detail the accuracy of the algorithm under various conditions at 4Hz and 10Hz data resolution. The initial speed profile data used in the simulation model was taken from an ATC site with a 40mph speed limit but additional results from model runs with speed profiles adjusted for a 30mph speed limit are also included. In each scenario the model results are averaged across five random seed runs with the performance in under-saturated (<S) and oversaturated (S) conditions reported separately. Each model maintained a representative distribution of vehicle lengths recorded from the ATC site.

Table 1. Results of the 4Hz and 10Hz algorithm accuracy under various conditions.

	40mph speed limit				30mph speed limit			
	4Hz		10Hz		4Hz		10Hz	
	<S	S	<S	S	<S	S	<S	S
Low LV% (5%)	86%	68%	96%	78%	85%	69%	97%	85%
Mid LV% (10%)	86%	65%	92%	78%	84%	68%	95%	82%
High LV% (20%)	87%	61%	94%	65%	83%	58%	93%	63%

Performance of the algorithm clearly deteriorates, more significantly for the 20% LV proportion (an unlikely scenario for the vast majority of urban links), as a result of the algorithm being based on the confidence of the previous vehicle speed estimate. This is exaggerated by classification at very low speed being based purely on the dominant vehicle class. 10Hz data, as would be expected, provides more accurate results than 4Hz in all conditions with substantial improvements in oversaturated conditions. The different speed profiles have little effect on algorithm performance although data resolution limitations are likely to affect results at significantly higher speeds.

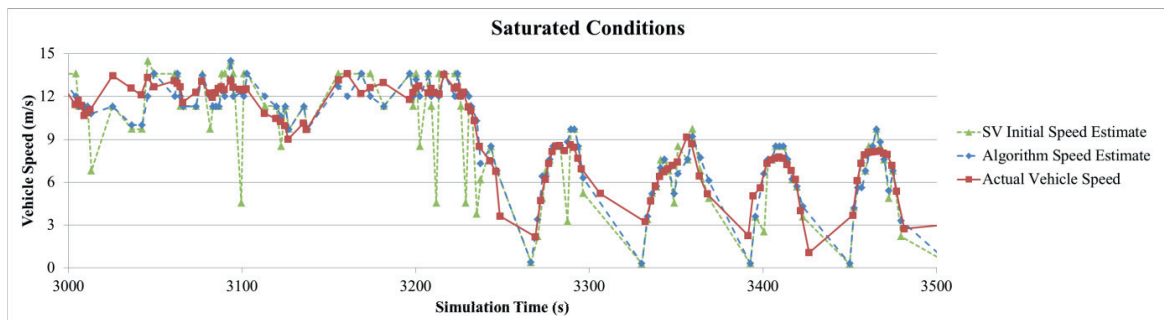


Fig. 6 shows an example plot of vehicle speeds estimated by the developed algorithm in saturated conditions compared to the actual vehicle speeds recorded from Vissim. Note that there is a slight discrepancy, more noticeable at slower speeds, between the time the vehicle speed was recorded by the algorithm and the vehicle record data. This is due to the frequency that the vehicle record database is updated by Vissim.

## 5. Conclusions

The reason for this research was to ascertain the level of classification accuracy that could be achieved for the purposes of creating a more sophisticated traffic model for signal optimisation. The 10Hz algorithm in particular shows promise for this application and given that only three parameters need calibrating, the SV and LV EVLs and average freeflow speed for the link on which the detector is sited, it can be considered a robust approach. The EVLs for the algorithm are easy to calibrate, being approximately the sum of the length of a standard car or HGV and the length of the detector, whilst the average freeflow speed can be calibrated from observations.

As a standalone algorithm this work could prove useful to traffic authorities for monitoring LV proportions on the urban road network although more investigation would be needed to prove algorithm performance on links with different characteristics and at different detector distances from a traffic signal stopline. The ‘instant’ nature of the speed estimation, as opposed to five minute aggregation, provides a quick response to drops in vehicle speed that could be used for incident detection purposes.

Not recorded in the results are the occasional instances of ‘missed’ vehicles where, at slow speeds, the detector does not see the gap between vehicles and as a result does not record or classify them. A future study could test a shorter detector length (i.e. 1m) to determine whether it would reduce this type of error significantly without adversely affecting accuracy at higher speeds as a result of measurement errors discussed previously.

The next step for establishing whether the developed algorithm performs well enough to create a more sophisticated traffic model that is sufficiently reliable is to test it within a simple optimisation process. A traffic model fed information from the algorithm can be compared to a model using ‘traditional’ information such as that used by SCOOT. The performance gains provided by more detailed classification data such as from additional roadside equipment or, as should be expected in future, cooperative technology could also be quantified.

## References

- Bonsall, P. et al. 2005. Modelling safety-related driving behaviour—impact of parameter values. *Transportation Research Part A: Policy and Practice*. **39**(5), pp.425-444.
- Coifman, B. 2001. Improved velocity estimation using single loop detectors. *Transportation Research Part A: Policy and Practice*. **35**(10), pp.863-880.
- Coifman, B. and Kim, S. 2009. Speed estimation and length based vehicle classification from freeway single-loop detectors. *Transportation Research Part C: Emerging Technologies*. **17**(4), pp.349-364.
- Highways Agency. 2007. *Guidelines for the Use of Microsimulation Software*.
- Kwon, J. et al. 2003. Estimation of Truck Traffic Volume from Single Loop Detectors with Lane-to-Lane Speed Correlation. *Transportation Research Record: Journal of the Transportation Research Board*. **1856**(-1), pp.106-117.
- Peck. 2012. *Chameleon Outstation User Handbook*. Steel, P.A., M.; Hodge, A.; Saini, O.; Leigh, A.; Clayton, D.A.; Fulton, L.; Pursey, R.; Ball, J.; Foote, M.; Bounds, D. p.86. [User Handbook].
- Robertson, D.I. and Bretherton, R.D. 1991. Optimizing networks of traffic signals in real time-the SCOOT method. *Vehicular Technology, IEEE Transactions on*. **40**(1), pp.11-15.
- RTEM. 2013. *SP4 Loop Detection Platform*. [Product Datasheet].
- Sun, C. and Ritchie, S. 1999. Individual Vehicle Speed Estimation Using Single Loop Inductive Waveforms. *Journal of Transportation Engineering*. **125**(6), pp.531-538.
- Wang, Y. and Nihan, N.L. 2003. Can single-loop detectors do the work of dual-loop detectors? *Journal of Transportation Engineering*. **129**(2), pp.169-176.
- Zhanfeng, J. et al. 2001. The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors. In: *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE, 2001*, pp.536-541.